



Year: 2020

Benefit-harm balance of fingolimod in patients with MS: A modelling study based on FREEDOMS

Spanu, Alessandra ; Aschmann, Hélène E ; Kesselring, Jürg ; Puhan, Milo A

Abstract: **BACKGROUND:** Fingolimod lowers the number of relapses in multiple sclerosis (MS) patients and slows down disease progression, but causes a broad spectrum of side effects. Our aim was to estimate the benefit-harm balance of fingolimod using individual patient data from FREEDOMS, a randomized controlled trial that compared two different dosages of fingolimod to placebo. **METHODS:** We modelled the health status of patients over two years on a scale ranging from 0 (worst health or death) to 100 (maximum health). The model considered Expanded Disability Status Scale measurements, relapses and adverse events. We compared the mean health status between arms, and the proportion of trial participants for whom health declined or improved compared to baseline by a predefined minimal important difference of 4.6 or more. **RESULTS:** The main analysis showed a net benefit for fingolimod 0.5mg compared to placebo, with an average health status difference over two years of 2.7 (95% CI 2.2 to 3.2). Patients on fingolimod 0.5mg were 0.53 (95% CI 0.40-0.72, $p < 0.001$) times less likely to have a relevant decline in health status compared to patients on placebo, corresponding to a number needed to treat of 8 to prevent one relevant decline in health status. All sensitivity analyses favoured fingolimod 0.5mg. **CONCLUSION** Although fingolimod's net benefit did not reach the clinical relevance on average, the decreased risk for a decline in health over two years may be relevant. This approach could be applied to other MS drugs and provide an objective evidence base for guideline recommendations.

DOI: <https://doi.org/10.1016/j.msard.2020.102464>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-191088>

Journal Article

Published Version

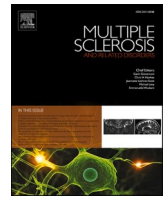


The following work is licensed under a Creative Commons: Attribution 4.0 International (CC BY 4.0) License.

Originally published at:

Spanu, Alessandra; Aschmann, Hélène E; Kesselring, Jürg; Puhan, Milo A (2020). Benefit-harm balance of fingolimod in patients with MS: A modelling study based on FREEDOMS. *Multiple Sclerosis and Related Disorders*, 46:102464.

DOI: <https://doi.org/10.1016/j.msard.2020.102464>



Benefit-harm balance of fingolimod in patients with MS: A modelling study based on FREEDOMS

Alessandra Spanu^a, Hélène E. Aschmann^a, Jürg Kesselring^b, Milo A. Puhan^{a,*}

^a Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Switzerland

^b RehaKliniken Valens, Zurich, Switzerland

ARTICLE INFO

Keywords:

Benefit-harm balance
Multiple sclerosis
Fingolimod
Health status
Adverse events
Expanded Disability status scale

ABSTRACT

Background: Fingolimod lowers the number of relapses in multiple sclerosis (MS) patients and slows down disease progression, but causes a broad spectrum of side effects. Our aim was to estimate the benefit-harm balance of fingolimod using individual patient data from FREEDOMS, a randomized controlled trial that compared two different dosages of fingolimod to placebo.

Methods: We modelled the health status of patients over two years on a scale ranging from 0 (worst health or death) to 100 (maximum health). The model considered Expanded Disability Status Scale measurements, relapses and adverse events. We compared the mean health status between arms, and the proportion of trial participants for whom health declined or improved compared to baseline by a predefined minimal important difference of 4.6 or more.

Results: The main analysis showed a net benefit for fingolimod 0.5mg compared to placebo, with an average health status difference over two years of 2.7 (95% CI 2.2 to 3.2). Patients on fingolimod 0.5mg were 0.53 (95% CI 0.40-0.72, $p < 0.001$) times less likely to have a relevant decline in health status compared to patients on placebo, corresponding to a number needed to treat of 8 to prevent one relevant decline in health status. All sensitivity analyses favoured fingolimod 0.5mg.

Conclusion: Although fingolimod's net benefit did not reach the clinical relevance on average, the decreased risk for a decline in health over two years may be relevant. This approach could be applied to other MS drugs and provide an objective evidence base for guideline recommendations.

1. Introduction

Fingolimod is an oral drug approved at the dosage of 0.5mg once daily for treating patients affected by relapsing remitting multiple sclerosis (RRMS), the most common form of multiple sclerosis (MS) (Goldenberg, 2012). A systematic review of randomized controlled trials (RCTs) found that fingolimod 0.5mg lowered the risk of relapses compared to placebo and interferon beta 1a (La Mantia et al., 2016). However, RCTs and post marketing surveillance reported adverse events (AEs) that raised concerns about the safety of fingolimod. The uncertainty around the benefit-harm balance led to different indications for fingolimod across health care systems. The NICE institute (NICE, 2012), for instance, recommends fingolimod only for highly active RRMS, while the FDA approved its usage in a broader range of MS forms (FDA, 2019).

Systematic reviews and health technology assessments usually consider benefits and harms separately and rarely use a systematic,

quantitative approach to combine all relevant outcomes in order to judge the benefit-harm balance (Puhan et al., 2013). If many benefit and harm outcomes are patient-important, it is challenging to judge the balance of benefits and harms without a systematic approach. In the absence of systematic, quantitative assessments, regulatory decisions and guideline recommendations can be discrepant (Yu et al., 2014; Yebo et al., 2019).

Instead, quantitative modelling can help clarify the benefit-harm balance of a treatment. Therefore, to clarify the benefit-harm balance of fingolimod, we performed a modelling study based on individual patient data (IPD) from FREEDOMS (FTY720 Research Evaluating Effects of Daily Oral Therapy in Multiple Sclerosis) (Kappos et al., 2010). This RCT compared daily doses of 0.5mg vs 1.25mg of fingolimod vs placebo in patients affected by RRMS. Analysing IPD has numerous advantages compared to aggregate data: it allows for incorporating follow-up assessments, relapses and adverse events as well as the

* Corresponding author: Milo A. Puhan, Hirschengraben 84, 8001 Zurich
E-mail address: miloalan.puhan@uzh.ch (M.A. Puhan).

<https://doi.org/10.1016/j.msard.2020.102464>

Received 18 May 2020; Received in revised form 12 August 2020; Accepted 22 August 2020

Available online 23 August 2020

2211-0348/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

sequence and co-occurrence of these events within individual patients.

2. Methods

2.1. Study design

We performed a quantitative benefit-harm modelling study based on IPD of the phase 3 trial FREEDOMS. Briefly, we first estimated the health status of each participant over the 24-months study period based on the Expanded Disability Status Scale (EDSS) (Kurtzke, 1983) score at baseline and at scheduled visits and additionally considered all relapses and AEs. We then compared the health status between participants of the three trial arms.

FREEDOMS compared two different dosages of fingolimod, 1.25mg or 0.5mg once daily (later referred to as fingolimod 1.25mg and fingolimod 0.5mg) to matching placebo. The trial population consisted of patients with RRMS who were between 18 to 55 years of age, had a baseline EDSS score between 0 to 5.5, and experienced at least one relapse in the year before enrolment, or two or more relapses in the two years before enrolment. Patients were eligible if they had prior treatment with interferons, glatiramer acetate and natalizumab (see Supplemental Data, Section 3) but with a washout period of at least three months before randomization.

We were granted access to IPD of FREEDOMS through the Clinical Study Data Request platform (CSDR, ClinicalStudyDataRequest.com) that offers access to trials released by a consortium of sponsors. Our application was approved by the platform Independent Research Board and a data sharing agreement (1933) was signed by the study sponsor Novartis and us. The sponsor of the trial had no influence on the study question, the methods, the analysis and the interpretation of the results, writing of the manuscript or the decision to publish.

2.2. Estimating the health status for individuals based on EDSS, relapses and adverse events

We modelled the health status of each individual in the trial on a scale from 100 (perfect health) to 0 (death, as assumed to represent worst health status), as commonly used in health economic analyses or burden of disease studies (Department of Information E and R, WHO G, 2017). In a first step, we estimated the health status based on the EDSS at baseline and at scheduled visits (see Supplemental Data, Fig. S4). In a second step, we reduced the health status if the individual experienced relapses or AEs.

We accounted for relapse severity (mild, moderate, severe, (Naldi et al., 2011; Ahmad and Taylor B, 2017) Supplemental Data, section 1.2.1) and duration, including whether patients were hospitalized and/or treated with systemic corticosteroids. We considered only relapses confirmed by the EDSS, but we had to attribute them a drop in health status based on the reported severity since EDSS scores measured for relapse assessments were not available in the dataset. Relapses were classified as mild, moderate, severe, very severe; if a relapse was coded as mild, the EDSS rise compared to the stable status was estimated at 1.7 (the central value for the range, see Table 1) and the corresponding health status drop was 9.1, assuming no treatment. The population under study did not experience very severe relapses. We considered all AEs, which were coded using preferred terms of the Medical Dictionary for Regulatory Activities (MedDRA). We also considered their severity (no symptoms, mild, moderate, severe), duration and therapeutic actions taken to treat them, which were all reported in the datasets. “Multiple sclerosis relapse” and “Multiple Sclerosis” AEs were excluded to avoid double counting of relapses. We decided to not incorporate Multiple Sclerosis Functional Composite (MSFC) assessments in our analysis, since they could lead to double counting of the functional status. Furthermore, we designed a method applicable to different studies and with the MSFC we would face a challenge for applicability of the methods.

2.3. Drops in health status due to relapses and adverse events

We decided a priori how much the health status dropped due to a relapse or an AE of a specific severity. We applied the same drops to all patients irrespective of group assignment. The health status of patients dropped the first day of a relapse or AE and then gradually recovered from the second day of the event (Supplemental Data, section 1.2.1 and 1.2.2). Preference surveys that elicit the importance of each specific event would be ideal to empirically determine the drop in health status (Aschmann et al.; Aschmann et al., 2019; Yebyo et al., 2018; Zhang et al., 2019), but we could not identify any such studies that covered the broad spectrum of AEs recorded in FREEDOMS. Instead, Table 1 and Table 2 show how we defined a drop in health status for each relapse and AE respectively.

Table 1 shows relapse drops used in the main analysis.

Table 2 shows Adverse Event categorisations and drops.

Drops increased with severity and depending on the action taken (details on how the drops were defined are in the Supplemental Data). The duration of an AE did not impact on the initial drop at day 1. AEs were coded in FREEDOMS as no symptoms, mild, moderate and severe, and we additionally defined four categories according to clinical judgement of how much impact an AE is likely to have (see AE Drops, Table 6, Supplemental Data).

To assign the drops we combined an algorithm with a manual change of single AE drop, when necessary. During this step it is possible that we created some discrepancies in drop assignment. However, this had no impact in our results, since AEs are coded equally in all arms and sensitivity analyses where we changed AEs drops showed not significant difference in comparison to the main findings (see sensitivity analyses for AEs). This process allowed us to account for different severities of the preferred terms within the same superior term (see AE Drops, Table 6, Supplemental Data, for superior and preferred term classification and drops).

We accounted in the model that more AEs could happen at the same time (Slankamenac et al., 2013). We performed a number of sensitivity analyses to assess how sensitive the benefit-harm balance of fingolimod was to our defined drops in health status (see Supplemental Data, Table 4 and Table 5). In two sensitivity analyses, we assigned larger or smaller drops to relapses compared to the main analysis. We assumed

Table 1
Drops in health status for relapses.

Outcome	Severity ^a	Treatment of relapse ^b	EDSS rise ^c	Health status drop
Relapse	Mild	None	1.7	9.1
		Systemic corticosteroids	-	9.6
		Hospitalization	-	22.0
		SC ^d + hospitalization	-	22.2
	Moderate	None	2.7	15.8
		Systemic corticosteroids	-	16.1
		Hospitalization	-	25.5
		SC + hospitalization	-	25.7
	Severe	None	3.5	21.9
		Systemic corticosteroids	-	22.1
		Hospitalization	-	29.7
		SC + hospitalization	-	29.8

^a Three categories based on Expanded Disability Status Scale;

^b Four categories, the latter being the combination of the previous two. (Further details on Action taken combination are provided in the Supplemental Data);

^c It shows the EDSS rise, corresponding to the health status drop compared to the stable status. EDSS scores are set according to relapse severity; the score was determined considering the central value for severity range of values (mild (0-2), moderate (2.5 -3), severe (>3))(11);

^d Systemic corticosteroids.

Table 2

Drops in health status for adverse events.

Severity of symptoms	Therapeutic action taken according to database	Category of adverse events according to potential impact on health status			
		Very small impact	Small impact	Moderate impact	Large impact
No symptoms	None	1.9	2.5	3.7	5.0
	Study drug dose adjusted	2.1	2.7	3.9	5.1
	Minimal therapy given	3.5	3.9	4.8	5.8
	Moderate therapy given	8.2	8.4	8.8	9.4
	Study drug dose suspended	20.1	20.2	20.3	20.6
Mild symptoms	(Prolonged) hospitalization	48.0	48.1	48.2	48.3
	None	3.8	5.0	7.5	10.0
	Study drug dose adjusted	3.9	5.1	7.6	10.1
	Minimal therapy given	4.8	5.8	8.1	10.4
	Moderate therapy given	8.8	9.4	11.0	12.8
Moderate symptoms	Study drug dose suspended	20.3	20.6	21.4	22.4
	(Prolonged) hospitalization	48.1	48.2	48.6	49.0
	None	7.5	10.0	15.0	20.0
	Study drug dose adjusted	7.6	10.1	15.0	20.0
	Minimal therapy given	8.1	10.4	15.3	20.2
Severe symptoms	Moderate therapy given	11.0	12.8	17.0	21.5
	Study drug dose suspended	21.4	22.4	25.0	28.3
	(Prolonged) hospitalization	48.6	49.0	50.3	52.0
	None	15.0	20.0	30.0	40.0
	Study drug dose adjusted	15.0	20.0	30.0	40.0
	Minimal therapy given	15.3	20.2	30.2	40.1
	Moderate therapy given	17.0	21.5	31.0	40.8
	Study drug dose suspended	25.0	28.3	36.1	44.7
	(Prolonged) hospitalization	50.3	52.0	56.6	62.5

This table lists the drops in health status for one adverse event and its severity and therapeutic consequence. Drops for combinations of adverse event are explained in the Supplemental Data.

1 Adverse events (preferred MedDRA terms) were separated into 4 categories based on clinical judgement about the impact the adverse event is likely to have on health status when mild and no action is taken. For example, nausea was classified as an adverse event with very small impact, bradycardia as an adverse event with small impact, viral bronchitis as an adverse event with moderate impact, and macular oedema as an adverse event with large impact on health status.

extreme drops for both sensitivity analyses in order to generate a range in which the difference in health status between groups most likely falls. We left all other model parameters constant for these sensitivity analyses.

Following the same logic, we conducted two sensitivity analyses to test how different drops for adverse events impacted on the results. In one sensitivity analysis, the drops increased less with severity than in the main analysis. In another sensitivity analysis, we did not consider mild AEs in order to explore the extent to which moderate and severe events alone impacted the benefit-harm balance of fingolimod.

Finally, we performed a sensitivity analysis with different assumptions to convert EDSS scores to a health status scale.

2.4. Censoring

In FREEDOMS, at two years, approximately 20% of patients with fingolimod 1.25mg and with placebo, and 10% of patients with fingolimod 0.5mg were censored. We used multiple imputation (25 imputed data sets) with “predictive mean matching” (package mice (Gaffert and Meinfelder, 2016)) to account for censoring.

2.5. Statistical analysis

We predefined a minimal important difference (MID) to judge the clinical relevance of the difference in health status between trial arms. We computed the MID as half the standard deviation of the health status at baseline (Norman et al., 2003), which equalled 4.6. A systematic review found that distribution-based approaches yielded more conservative, but similar values for the MID compared to anchor-based approaches (Jayadevappa et al., 2017). We did not find any survey which included all our outcomes, however in a study using anchor-based approach the MID was set at 1 for the EDSS range of 0-5.5 (Costelloe et al., 2007), which is identical to the range of baseline EDSS in FREEDOMS. Taking into account the non-linearity of the conversion from EDSS to the health status scale, this corresponds to a MID of about 8 on

the health status scale, which is larger than the distribution-based MID of 4.6. We preferred to use the distribution-based method, which is calculated directly on our sample; in addition, our main conclusions would not change using a MID of 8.

We calculated the difference of the mean health status between the groups. We defined that if the 95% confidence interval of this difference did not cross +4.6 or -4.6, the average health gain or loss would not be clinically relevant. Furthermore, we determined the proportion of patients who had a relevant improvement or decline (by 4.6 or more) compared to baseline after two years. We calculated the risk ratio between treatment groups for a relevant improvement or decline, respectively, and the number needed to treat.

All analyses were performed using R version 3.4.3 on the CTRD platform.

2.6. Data availability

The data that support the findings of this study are publicly available on the CSDR platform at request from the study sponsor Novartis.

3. Results

3.1. Main analysis

Fig. 1 shows the average health status in the three treatment groups over two years. The health status at baseline was around 85 for all three groups, which is, as expected, lower than the health status of a general population of similar age (which has an average health of 89 at age 35-45 (Fryback et al., 2007)).

Patients on fingolimod 0.5mg had the best health status over the entire study period. There were few relapses over the entire study period and after the first quarter of the year, adverse events were not infrequent but mostly of low to moderate severity. As a consequence, health status was stable across the two years of the trial.

In contrast we observed that the health status of patients in placebo

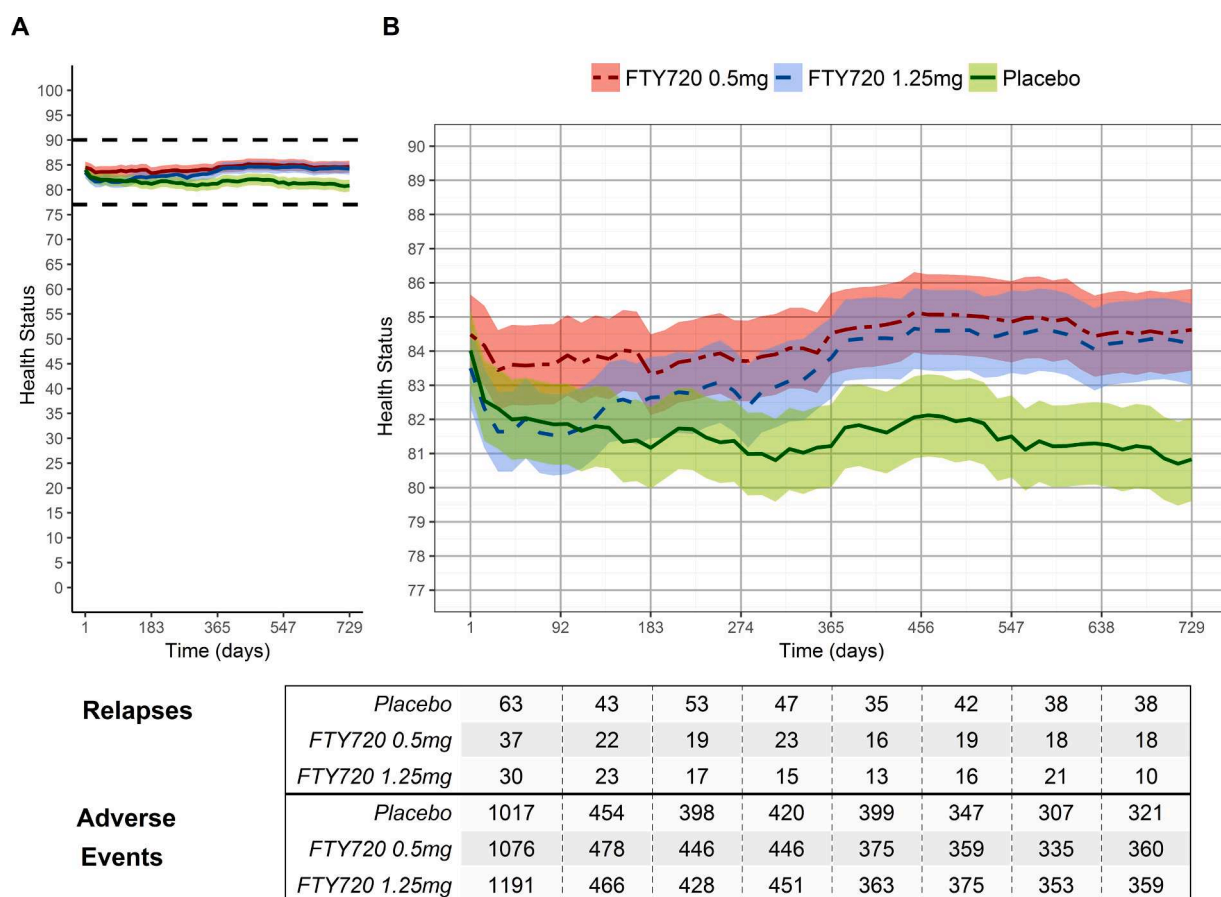


Fig. 1. Mean health status of patients with fingolimod 0.5mg, fingolimod 1.25mg and placebo over two years.

The figure shows the mean and 95% confidence interval of the mean health status in each group.

Panel A represents a full health status scale (0 to 100); Panel B shows a partial health status scale (y axis, from 77 to 90) and health status curves for the three treatment groups; below two tables show the total relapse numbers during the corresponding time periods, covering 91 days each (x axis) and the total adverse event numbers during the same corresponding time periods.

and fingolimod 1.25mg arms deteriorated within the first month, although in a different way. As an explanation for these deteriorations, Fig. 1 shows the number of relapses and adverse events for each quarter

year. Patients with placebo had almost twice the number of relapses compared to patients with fingolimod 0.5mg or 1.25mg in the first three months, and also substantially more relapses during the rest of the two

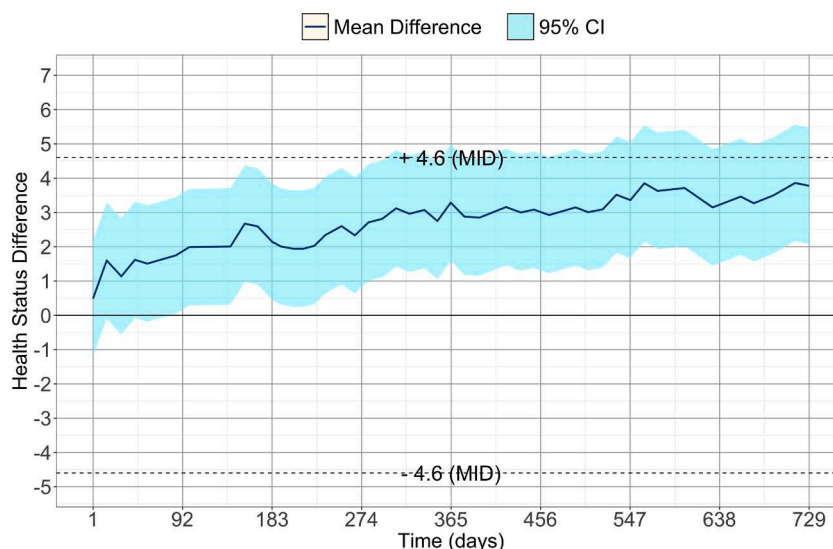


Fig. 2. Difference in mean health status between patients with fingolimod 0.5mg and placebo in the FREEDOMS trial.

If the difference is positive, patients on fingolimod had a better health status on average than those on placebo.

± 4.6 = Minimal Important Difference (MID).

years. As the study time progressed, we observed that these patients experienced a decreasing number of relapses over time. We observed a similar effect in the other two arms.

Patients included in the trial consisted of both naive and previously treated patients, with a washout period of three months prior to randomization. For the placebo group we found that among those with no relapses in the first three months, 36% of patients were previously treated, with one relapse 45% and with two or more relapses 80% were previously treated, respectively. In the fingolimod 0.5mg group, the corresponding proportions were 57%, 51% and no one with two or more relapses (see Supplemental Data, Section 3).

Throughout the two years, patients with placebo experienced more relapses than patients with fingolimod, which explains the constant decline of the health status. Health status of patients with fingolimod 1.25mg also deteriorated in the first 92 days of the trial, but this is likely due to the frequent and often severe adverse effects.

3.2. Difference in health status between fingolimod 0.5mg and placebo

Fig. 2 shows the difference in mean health status between patients on fingolimod 0.5mg and on placebo. The mean difference across two years was + 2.7 (95% CI 2.2 to 3.2). Thus the difference did not reach the MID of 4.6 points but was close to the MID in the last months of the trial (for results of fingolimod 1.25mg vs placebo, see Supplemental Data). At two years, the difference was 3.8 (95% CI 2.1 to 5.6).

We found that 27.0% (90/333) in the placebo group and 14.6% (54/369) in the fingolimod 0.5 mg experienced a relevant decline in health status (i.e. MID of 4.6 points) over two years compared to baseline. This corresponds to a risk ratio of 0.54 (95% CI 0.40-0.73) and a number needed to treat to prevent one relevant decline in health status of 8. 9.0% (30/333) in the placebo and 13.0 % (48/369) in the fingolimod 0.5mg group experienced a relevant improvement in health status over two years (risk ratio of 1.45, 95% CI 0.98-2.22, $p=0.12$), corresponding to a number needed to treat of 25. Where not specified otherwise, p values for previously mentioned CI were < 0.01

3.3. Sensitivity analyses with large and small drops for relapses

Different drops for relapses showed very similar results to those of the main analysis (Fig. 3). In the sensitivity analysis where we assigned a smaller drop to relapses in comparison to the main analysis (Fig. 3, *Relapse small drops*), the benefit-harm balance still favoured fingolimod 0.5mg, even though the difference over the entire study period (2.5, 95% CI 1.8 to 3.3) and towards the end of the trial (3.5 95% CI 1.9 to 5.1) was smaller. The Sensitivity Analysis with larger drops on relapses (Fig. 3, *Relapse large drops*) showed an average difference in health status

over two years of 3.6 (95% CI 2.6 to 4.6), thus clinical relevance cannot be rejected with statistical significance in this sensitivity analysis with extreme weight on relapses. This result represents the most favourable benefit-harm balance fingolimod 0.5mg obtained across all analyses. The sensitivity analyses demonstrated that the benefit-harm balance of fingolimod 0.5mg is only moderately sensitive to how relapses and adverse events are weighted.

3.4. Sensitivity analyses with different drops for Adverse Events

In a sensitivity analysis in which we assigned drops for AEs that increased less with severity (Fig. 3, *Adverse events small drops*), and a sensitivity in which we did not take into account mild AEs (Fig. 3, *Adverse events excl. mild*), respectively, there were no relevant differences compared to the main results (2.7, 95% CI 2.0 to 3.5, and 2.7, 95% CI 1.9 to 3.4 over two years, respectively). When we considered the proportion of patients exceeding the MID, the treatment was superior over the entire study duration (RR 1.4, 95% CI 1.0 to 2.0, $p = 0.05$).

3.5. Sensitivity analysis with different conversion of EDSS to health status scale

With different assumptions to convert EDSS scores to health status (Tworok et al., 2010) (Fig. 3, *Modified EDSS conversion*) and setting at 9 the worst possible value/death (Orme et al., 2007), the average difference in health status over two years was 3.4 (95% CI 2.4 to 4.3). This should be considered as an extreme variation of how EDSS scores can be converted to a health status scale, however results are in agreement with the main analysis.

P values are <0.01 in all sensitivity analyses.

4. Discussion

To our knowledge, a benefit-harm balance modelling study for fingolimod 0.5mg based on IPD from FREEDOMS was not been performed before. Our results showed that the benefit-harm balance favoured fingolimod 0.5mg over placebo with a difference that is statistically significant, with a number needed to treat of fingolimod 0.5mg to prevent one relevant decline in health status over two years of 8. All sensitivity analyses showed similar results favouring fingolimod 0.5mg irrespective of how the drops in health status due to relapses and adverse events were defined.

4.1. Clinical importance

Our findings imply that the benefits of fingolimod 0.5mg, i.e.

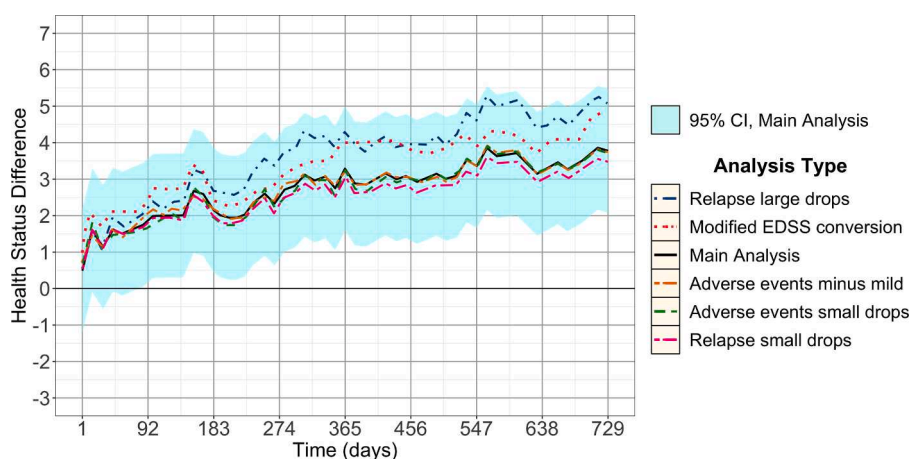


Fig. 3. Sensitivity analyses summary.

The main analysis and all sensitivity analyses are shown. The error band represents the 95% CI of main analysis difference.

reducing the relapse rate, outweigh the AEs, and that fingolimod 0.5mg stabilizes the patients' health status over time. Side effects were more frequent at the beginning of the study and decreased over time. The mean difference between the fingolimod 0.5mg and placebo groups was below the MID over the course of two years but almost reached the MID at the end of the study period. Interpreting the mean difference between groups with respect to the MID is conservative and may underestimate the net benefit of treatments. Therefore, we also determined the proportion of patients with a relevant decline in health status over two years in order to calculate the risk reduction of such a decline with fingolimod 0.5mg. The number needed to treat of fingolimod 0.5mg to prevent one relevant decline in health status over two years was 8 and indicates an absolute benefit. Whether this benefit is acceptable on a societal level depends on the cost effectiveness and cost impact associated with it, along with the context and resources of a particular country.

Concerning AEs, qualitative recommendations for monitoring patients on fingolimod 0.5mg treatment, particularly for cardiovascular events, infections and macular oedema (Oh and O'Connor, 2013), were issued as well as measures to minimize the risk of rare brain infections (FDA, 2019; Anton et al., 2017; European Medicines Agency 2013; FDA, 2012; Gilenya, 2012). Safety recommendations were based on case reports, AEs collected in RCTs and signals from post marketing surveillance. Our findings suggested that the safety profile of fingolimod 0.5mg is clearly better than fingolimod 1.25mg (a dosage that was not approved). Comparing fingolimod 0.5mg with placebo, the small number of severe cardiovascular events and infections had little impact on the benefit-harm balance. Progressive multifocal leukoencephalopathy (PML), a severe adverse event observed with many other disease-modifying MS drugs and associated with fingolimod 0.5mg intake (FDA), was never observed during FREEDOMS.

4.2. Preference sensitivity of benefit-harm balance

Our sensitivity analyses demonstrated that the model's results are stable, even with extreme assumptions for how much relapses decrease the health status. Since these extreme assumptions did not change the results meaningfully, it is unlikely that additional empirical evidence on average patient preferences would change the estimated benefit-harm balance of fingolimod 0.5mg on a population level. However, patient preferences can vary (Yebo et al., 2019; Aschmann et al.; Yu et al., 2015). Preferences could differ across patients depending on the course of MS, sex, age and other factors, and the individual benefit-harm balance may depend on the individual's preferences. For example, it is likely that not all patients would consider a number needed to treat of 8 to avoid a decline of 4.6 or more over two years as clinically relevant. This underscores the importance for informed and shared decision making since the overall results from RCTs and analyses such as ours do not generalize to individual patients and their preferences.

4.3. Strengths of the study

We modelled the benefit-harm balance of fingolimod 0.5mg based on IPD, which allowed considering all relapses and AE events and their sequence over time. We performed several sensitivity analyses, which showed that the benefit-harm balance is stable and variations in average preferences had little to no impact. In addition, we observed a high number of relapses early in the study in the placebo arm, that was followed by a decreasing relapse rate both of which were not described when FREEDOMS was published. This effect was smaller in fingolimod 0.5mg arm, which reached the peak efficacy over time.

4.4. Weaknesses of the study

The duration of FREEDOMS was likely too short to identify rare AEs such as PML. We do not know how the benefit-harm balance of fingolimod 0.5mg looks like beyond 2 years of treatment. Due to the relatively

small sample size of the trial, we could not perform subgroup analyses, and therefore we do not know if some groups may benefit more or less from fingolimod 0.5mg. We observed a greater number of AEs at the beginning than in the second year of the study, even in the placebo group. The strong decrease in AEs over time in all arms could be a consequence of a tolerance effect, and over-reporting at the beginning or under-reporting later on in the study.

5. Conclusion

Our results showed that the benefit-harm balance favoured fingolimod 0.5 mg over placebo in a statistically significant manner. Fingolimod's preventive effect reducing the risk of a decline in health status is likely to be clinically meaningful on a population level, despite the fact that the difference in health status did not reach the MID on average. This effect could, however, vary according to individual preferences depending on how a patient perceives the importance of relapses and adverse events, potentially changing the benefit-harm balance across patients.

Our benefit-harm balance modelling study could pave the way for similar analyses of other MS drugs where the benefit-harm balance is debated and, thereby, provide an important and objective evidence base for guideline recommendations.

Study funding

AS received PhD funding by the University of Zurich. HEA received a PhD fellowship from the Béatrice Ederer-Weber foundation.

CRediT authorship contribution statement

Alessandra Spanu: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Hélène E. Aschmann:** Methodology, Validation, Formal analysis, Writing - review & editing. **Jürg Kesselring:** Validation, Writing - review & editing, Visualization. **Milo A. Puhan:** Conceptualization, Methodology, Validation, Formal analysis, Resources, Writing - review & editing, Visualization, Supervision, Project administration.

Declaration of Competing Interests

Jürg Kesselring was a member of the Data Safety Monitoring Board of Fingolimod studies from 2005-2017.

The other authors report no disclosures.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.msard.2020.102464](https://doi.org/10.1016/j.msard.2020.102464).

References

- Ahmad H., Taylor B V., van der Mei I, Colman S, O'Leary BA, Breslin M, et al. The impact of multiple sclerosis severity on health state utility values: Evidence from Australia. *Mult. Scler.* 2017 Jul;23(8):1157–66.
- Anton, R., Haas, M., Arlett, P., Weise, M., Balabanov, P., Mazzaglia, G., et al., 2017 Aug. Drug-induced progressive multifocal leukoencephalopathy in multiple sclerosis: European regulators' perspective. *Clin. Pharmacol. Ther.* 102 (2), 283–289.
- Aschmann, H.E., Boyd, C.M., Robbins, C.W., Chan W. V., Mularski, R.A., Bennett, W.L., et al., 2019. Informing patient-centered care through stakeholder engagement and highly stratified quantitative benefit-harm assessments. *Value Heal.*
- Aschmann H.E., Puhan M.A., Robbins C.W., Bayliss E.A., Chan W V., Mularski R.A., et al. Outcome preferences of older people with multiple chronic conditions and hypertension: a cross-sectional survey using best-worst scaling. *HQLO*, accept pending minor revis.
- Costelloe, L., O'Rourke, K., Kearney, H., McGuigan, C., Gribbin, L., Duggan, M., et al., 2007 Aug. The patient knows best: significant change in the physical component of the Multiple Sclerosis Impact Scale (MSIS-29 physical). *J. Neurol. Neurosurg. Psychiatry* 78 (8), 841–844.

- Department of Information E and R, WHO G. WHO methods and data sources for global burden of disease estimates 2000-2015. 2017.
- European Medicines Agency. Annual report on Eudravigilance for the European parliament, the council and the commission. [Internet]. 2013. https://www.ema.europa.eu/en/documents/report/2013-annual-report-eudravigilance-european-parliament-council-commission_en.pdf.
- FDA Tysabri (natalizumab): drug safety communication - new risk factor for Progressive Multifocal Leukoencephalopathy (PML). [Internet]. 2012. <http://www.fda.gov/Safety/MedWatch/SafetyInformation/SafetyAlertsforHumanMedicalProducts/ucm288602.htm>.
- FDA. Medication guide gilenya® [Internet]. 2019. https://www.accessdata.fda.gov/drugsatfda_docs/label/2019/022527s029s030lbl.pdf#page=26.
- FDA. FDA drug safety communication: FDA investigating rare brain infection in patient taking Gilenya (fingolimod) [Internet]. <http://wayback.archive-it.org/7993/20170112031628/http://www.fda.gov/Drugs/DrugSafety/ucm366529.htm>.
- Fryback, D.G., Dunham, N.C., Palta, M., Hanmer, J., Buechner, J., Cherepanov, D., et al., 2007 Dec. US norms for six generic health-related quality-of-life indexes from the National Health Measurement study. *Med. Care* 45 (12), 1162–1170.
- Gaffert, P., Meinfelder, F.B.V., 2016. Towards an MI-proper predictive mean matching. In: *Conf Proc* [Internet]. <https://www.semanticscholar.org/paper/Towards-an-MI-proper-Predictive-Mean-Matching-Gaert-Meinfelder/5f13dd8d367c9b0ebc127c24487aa06773a2f00f>.
- FDA. Gilenya(fingolimod): drug safety communication - safety review of a reported death after the first dose. [Internet]. 2012. <https://www.fda.gov/drugs/drug-safety-and-availability/fda-drug-safety-communication-safety-review-reported-death-after-first-dose-multiple-sclerosis-drug>.
- Goldenberg, M.M., 2012 Mar. *Mult. Scler. Rev. P T* 37 (3), 175–184.
- Jayadevappa, R., Cook, R., Chhatre, S., 2017 Sep. Minimal important difference to infer changes in health-related quality of life-a systematic review. *J. Clin. Epidemiol.* 89, 188–198.
- Kappos, L., Radue, E.-W., O'Connor, P., Polman, C., Hohlfeld, R., Calabresi, P., et al., 2010 Feb. A placebo-controlled trial of oral fingolimod in relapsing multiple sclerosis. *N. Engl. J. Med.* 362 (5), 387–401.
- Kurtzke, J.F., 1983 Nov. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 33 (11), 1444–1452.
- La Mantia, L., Tramaccere, I., Firwana, B., Pacchetti, L., Palumbo, R., Filippini, G., 2016 Apr 19. Fingolimod for relapsing-remitting multiple sclerosis. *Cochrane Database Syst. Rev.* [Internet]. <https://doi.org/10.1002/14651858.CD009371.pub2> [cited 2019 Jul 8];(4).
- Naldi, P., Collimedaglia, L., Vecchio, D., Rosso, M.G., Perl, F., Stecco, A., et al., 2011. Predictors of attack severity and duration in multiple sclerosis: a prospective study. *Open Neurol. J.* 5, 75–82.
- NICE. Fingolimod for the treatment of highly active relapsing-remitting multiple sclerosis [Internet]. 2012. <https://www.nice.org.uk/guidance/ta254/chapter/4-Consideration-of-the-evidence>.
- Norman, G.R., Sloan, J.A., Wyrwich, K.W., 2003 May. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med. Care* 41 (5), 582–592.
- Oh, J., O'Connor, P.W., 2013 Aug. Safety, tolerability, and efficacy of oral therapies for relapsing-remitting multiple sclerosis. *CNS Drugs* 27 (8), 591–609.
- Orme, M., Kerrigan, J., Tyas, D., Russell, N., Nixon, R., 2007. The effect of disease, functional status, and relapses on the utility of people with multiple sclerosis in the UK. *Value Heal. J. Int. Soc. Pharmacoecon. Outcomes Res.* 10 (1), 54–60.
- Puhan M.A., Singh S., Weiss C.O., Varadhan R., Sharma R., Boyd C.M. Evaluation of the benefits and harms of aspirin for primary prevention of cardiovascular events: a comparison of quantitative approaches. Rockville (MD); 2013.
- Slankamenac, K., Graf, R., Barkun, J., Puhan, M.A., Clavien, P.A., 2013. The comprehensive complication index: a novel continuous scale to measure surgical morbidity. *Ann. Surg.* 258 (1), 1–7.
- Two, S., Wiesmeth, S., Spindler, M., Wirtz, M., Schipper, S., Pöhlau, D., et al., 2010 Jun. Disability status and quality of life in multiple sclerosis: non-linearity of the Expanded Disability Status Scale (EDSS). *Health Qual. Life Outcomes* 8, 55.
- Yebo, H.G., Aschmann, H.E., Puhan, M.A., 2019. Finding the balance between benefits and harms when using statins for primary prevention of cardiovascular disease: a modeling study. *Ann. Intern. Med.* [Internet] 170 (1), 1–10. <https://doi.org/10.7326/M18-1279>.
- Yebo, H.G., Aschmann, H.E., Yu, T.P.M., 2018. Should statin guidelines consider patient preferences? Eliciting preferences of benefit and harm outcomes of statins for primary prevention of cardiovascular disease in the sub-Saharan African and European contexts. *BMC Cardiovasc. Disord.* 18, 97.
- Yu, T., Holbrook, J.T., Thorne, J.E., Flynn, T.N., Van Natta, M.L., Puhan, M.A., 2015. Outcome preferences in patients with non infectious uveitis: results of a best–worst scaling study. *Invest. Ophthalmol. Vis. Sci.* [Internet] 56 (11), 6864–6872. <https://doi.org/10.1167/iov.15-16705>. Available from.
- Yu, T., Fain, K., Boyd, C.M., Singh, S., Weiss, C.O., Li, T., et al., 2014 Jul. Benefits and harms of roflumilast in moderate to severe COPD. *Thorax* 69 (7), 616–622.
- Zhang, Y., Alonso-Coello, P., Guyatt, G.H., Yepes-Nunez, J.J., Akl, E.A., Hazlewood, G., et al., 2019 Jul. GRADE guidelines: 19. Assessing the certainty of evidence in the importance of outcomes or values and preferences-Risk of bias and indirectness. *J. Clin. Epidemiol.* 111, 94–104.